Q.2 a. What are main objectives of AI research?

Answer: Refer Page Number 13

b. Differentiate between symbolic and non-symbolic representation.

Answer: Refer Page Number 5

Q.3 a. For each pair of atomic sentences given below, give the most general unifier if it exists.

(i) P(A,B,B), P(x,y,z)
(ii) Q(Y,G(A,B)), Q(G(x,x),y)
(iii) Older(Father(y),y), Older(father(x), John)
(iv) Knows(father(y),y), Knows(x, x)

Answer:

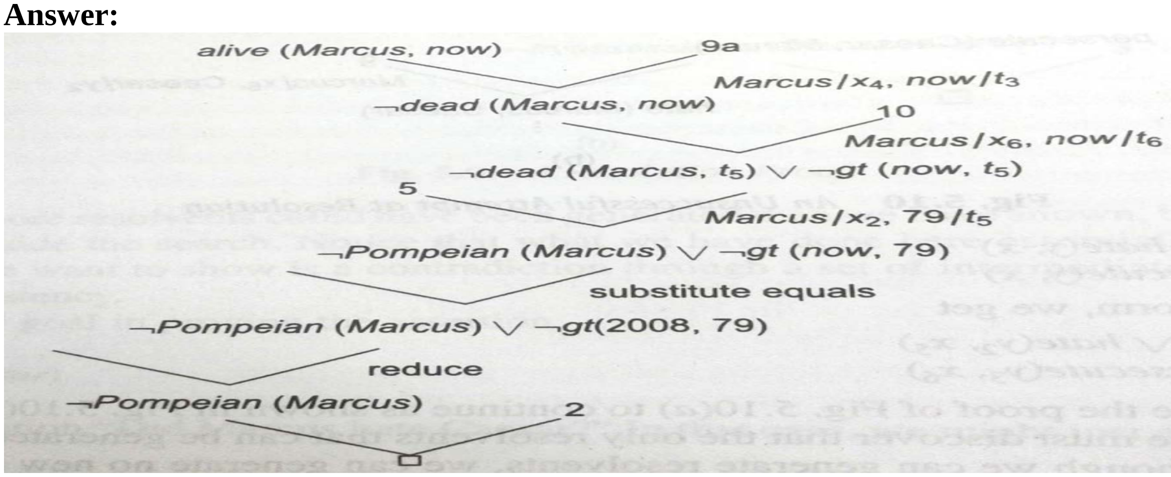
i. {x/A,y/B,z/B} (or Some permutation of this)
ii. No unifier (x can not bind to both A and B)
iii. {y/John,x/John}
iv. No unifier

b. For the following axioms :

1. **man(Marcus)**
2. **pompeian(Marcus)**
3. **horn(Marcus,40)**
4. **$\neg$man(x_1)V mortal(x_1)**
5. **$\neg$pompeian (x_2)Vdied(x_2,79)**
6. **erupted(volcano,79)**
7. **$\neg$morta(x_3)V$\neg$born(x_3,t_1)V$\neg$gt(t_2-t_1,150)Vdead(x_3,t_2)**
8. **now=2028**
9. **$\neg$alive(x_4,t_3)V$\neg$dead(x_4,t_3)**
10. **dead(x_5,t_4)Valive(x_5,t_4)**
11. **$\neg$dead(x_6,t_5)V $\neg$ gt(x_6,t_5)V dead(x_6,t_6)**

Prove $\neg$alive(Marcus,now).

Answer:



c. When is a WFF said to be :-
(i) Valid **(ii) Satisfiable**

Answer: Refer Page No 34

Q.4 a. Write four properties that a good system must possess for the representation of knowledge. Also write issues in knowledge representation.

Answer:

1. Representational Adequacy-Ability to represent all kinds of knowledge that are needed in that domain

2. Inferential Adequacy-Ability to manipulate the representational structures in such a way as to derive new structures corresponding to new knowledge inferred from old.

3. Inferential Efficiency-the ability to incorporate into the knowledge structure the additional information that can be used to focus the attention of the inference mechanism in the most promising directions.

4. Acquisitional Efficiency-the ability to acquire new information easily.

Issues

Important Attributes

-- Are there any attributes that occur in many different types of problem?

Relationships

-- What about the relationship between the attributes of an object, such as, inverses, existence, techniques for reasoning about values and single valued attributes.

Granularity

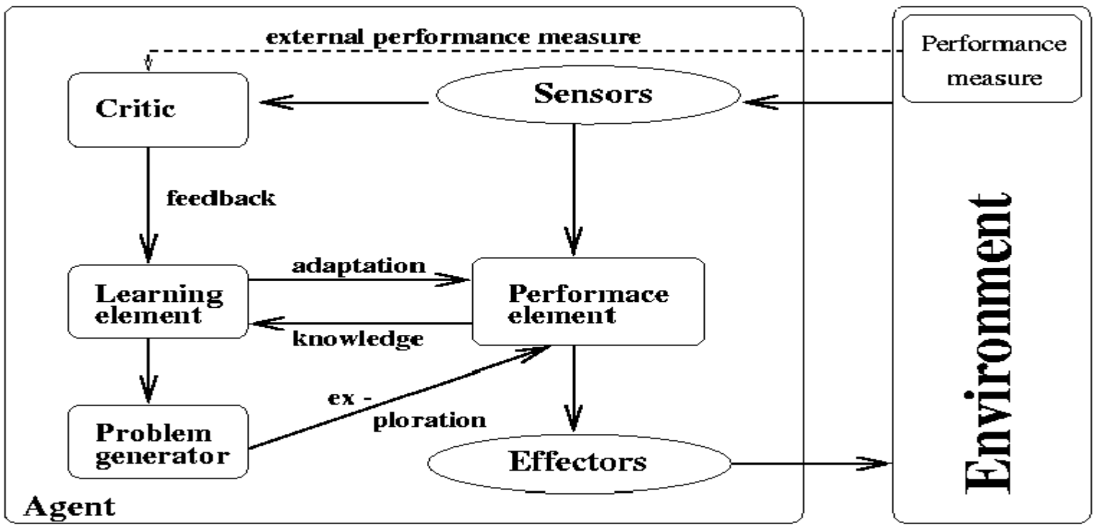
-- At what level should the knowledge be represented and what are the primitives. Choosing the Granularity of Representation Primitives are fundamental concepts such as holding, seeing, playing and as English is a very rich language with over half a million words it is clear we will find difficulty in deciding upon which words to choose as our primitives in a series of situations.

b. Explain procedure for knowledge acquisition.

Answer: Refer Page Number 53

Q.5 a. What is Hybrid Representation system? Explain any one such representation system.

Answer:
A hybrid KR system is an implementation of a hybrid KR formalism consisting of two or more different sub formalisms. These sub formalism should be integrated through (i) a representational theory, which explains what knowledge is to be-represented by what formalism, and (ii) a common semantics for the overall formalism, explaining in a semantic sound manner the relationship between expressions of different sub formalisms. The generalized architecture for a hybrid system is given in Fig



In general these systems consist of two different kinds of knowledge: The terminological knowledge, consisting of a set of concepts and roles defining a terminology, and the assertional knowledge, consisting of some logical formalism suited to represent general assertions

b. Explain Dempster and Shafer's theory of evidences in detail.

Answer:
The Dempster–Shafer theory (DST**)** is a mathematical theory of evidence. It allows one to combine evidence from different sources and arrive at a degree of belief (represented by a belief function) that takes into account all the available evidence. The theory was first developed by Arthur P. Dempster and Glenn Shafer.
Dempster–Shafer theory is a generalization of the Bayesian theory of subjective probability; whereas the latter requires probabilities for each question of interest, belief functions base degrees of belief (or confidence, or trust) for one question on the probabilities for a related question. These degrees of belief may or may not have the mathematical properties of probabilities; how much they differ depends on how closely the two questions are related.[5] Put another way, it is a way of representing epistemic plausibilities but it can yield answers that contradict those arrived at using probability theory.
Often used as a method of sensor fusion, Dempster–Shafer theory is based on two ideas: obtaining degrees of belief for one question from subjective probabilities for a related

question, and Dempster's rule[6] for combining such degrees of belief when they are based on independent items of evidence. In essence, the degree of belief in a proposition depends primarily upon the number of answers (to the related questions) containing the proposition, and the subjective probability of each answer. Also contributing are the rules of combination that reflect general assumptions about the data.
In this formalism a degree of belief (also referred to as a mass) is represented as a belief function rather than a Bayesian probability distribution. Probability values are assigned to *sets* of possibilities rather than single events: their appeal rests on the fact they naturally encode evidence in favor of propositions.
Dempster–Shafer theory assigns its masses to all of the non-empty subsets of the entities that compose a system
Formal definition
Let X be the *universal set*: the set representing all possible states of a system under consideration. The power set

$$2^X$$

is the set of all subsets of X, including the empty set $\varnothing$. For example, if:

$$X = \{a, b\}$$

then

$$2^X = \{\varnothing, \{a\}, \{b\}, X\}.$$

The elements of the power set can be taken to represent propositions concerning the actual state of the system, by containing all and only the states in which the proposition is true.
The theory of evidence assigns a belief mass to each element of the power set. Formally, a function

$$m : 2^X \rightarrow [0, 1]$$

is called a *basic belief assignment* (BBA), when it has two properties. First, the mass of the empty set is zero:

$$m(\varnothing) = 0.$$

Second, the masses of the remaining members of the power set add up to a total of 1:

$$\sum_{A \in 2^X} m(A) = 1$$

The mass $m(A)$ of A, a given member of the power set, expresses the proportion of all relevant and available evidence that supports the claim that the actual state belongs to A but to no particular subset of A. The value of $m(A)$ pertains *only* to the set A and makes no additional claims about any subsets of A, each of which have, by definition, their own mass.
From the mass assignments, the upper and lower bounds of a probability interval can be defined. This interval contains the precise probability of a set of interest (in the classical sense), and is bounded by two non-additive continuous measures called **belief** (or **support**) and **plausibility**:

$$\operatorname{bel}(A) \le P(A) \le \operatorname{pl}(A).$$

The belief $\operatorname{bel}(A)$ for a set A is defined as the sum of all the masses of subsets of the set of interest:

$$\operatorname{bel}(A) = \sum_{B \mid B \subseteq A} m(B).$$

The plausibility pl(*A*) is the sum of all the masses of the sets *B* that intersect the set of interest *A*:

$$\mathrm{pl}(A) = \sum_{B \mid B \cap A \neq \varnothing} m(B).$$

The two measures are related to each other as follows:

$$\mathrm{pl}(A) = 1 - \mathrm{bel}(\overline{A}).$$

And conversely, for finite *A*, given the belief measure bel(*B*) for all subsets *B* of *A*, we can find the masses m(*A*) with the following inverse function:

$$m(A) = \sum_{B \mid B \subseteq A} (-1)^{|A-B|}\,\mathrm{bel}(B)$$

where $|A - B|$ is the difference of the cardinalities of the two sets.
It follows from the last two equations that, for a finite set *X*, you need know only one of the three (mass, belief, or plausibility) to deduce the other two; though you may need to know the values for many sets in order to calculate one of the other values for a particular set. In the case of an infinite *X*, there can be well-defined belief and plausibility functions but no well-defined mass function.

Q.6 a. Describe Depth First Search algorithm and illustrate it with an example. Write its drawback also.

Answer:

Depth First Search

Depth-first: First, check all nodes in the same branch. In other words, visit each neighbour of the most recently visited node.

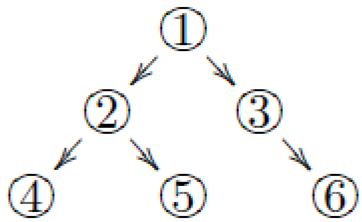
Depth first search (DFS) traverses a graph by going as deeply as possible before backtracking. It is surprisingly rich with potential for other algorithms. It also returns a search tree. It does not return the level of each node, but can return a numbering of the nodes in the order that they were visited. We first show a depth first search skeleton and define the different kinds classes of edges. Then we show how to augment the skeleton to solve two very basic algorithms: topological sorting, connected components. Each of leverages the power of DFS at a different location in the skeleton. We conclude with a sophisticated use of DFS that finds strongly connected components of a directed graph. You may recall that in month 0 we discussed a method in linear algebra using matrix multiplication that solved this algorithm in O(n^3). Our method will work in O (n+e)

Depth First Search Skeleton

```
DFS(G, s)
Mark s visited; Dfsnum[s] = count; count++; //count is a global counter initialized to 1.
/* Process s – previsit stage */
Recursive Loop:  For every y adjacent to s do
                        If y is unvisited then {DFS(G, y); parent[y] = x;} else…
```

/* process edges {s,y}*/;
/* Process s – postvisit stage */
Mark s finished

Example
Consider the following graph:



Starting at root node 1, give the order in which the nodes will be visited by the breadth-first and depth-first algorithms.
Answer:
Depth-first: 1 - 2 - 4 - 5 - 3 – 6
Drawback
The drawback of depth-first search is that it can make a wrong choice and get stuck going down a very long (or even infinite) path when a different choice would lead to a solution near the root of the search tree, hence, depth-first search is not optimal.

b. Explain heuristic search techniques briefly. Describe how it is applied in branch-and-bound search procedure.

Answer: Refer Page No 134

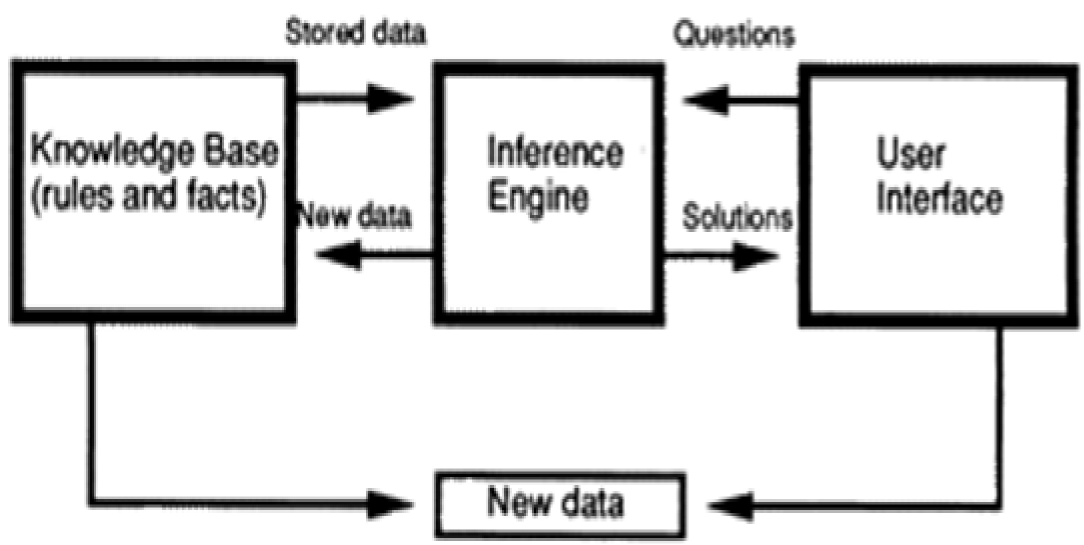
Q.7 a. Illustrate expert system with the help of a neat diagram and explain all components.

Answer:
An expert system is a system that employs human knowledge captured in a computer to solve problems that ordinarily require human expertise.(Turban)
A computer program that emulates the behaviour of human experts who are solving real-world problems associated with a particular domain of knowledge. (Pigford & Braur)
Solve simple problems easily.
Ask appropriate questions (based on external stimuli - sight, sound etc).
Reformulate questions to obtain answers.
Explain why they asked the question.
Explain why conclusion reached.
Judge the reliability of their own conclusions.
Talk easily with other experts in their field.
Learn from experience.
Reason on many levels and use a variety of tools such as heuristics, mathematical
Models and detailed simulations.

Transfer knowledge from one domain to another.
Use their knowledge efficiently

Major components of expert systems
(1.) The User Interface
The user interface is the means of communication between a user and the expert systems problem-solving processes. A good expert system is not very useful unless it has an effective interface. It has to be able to accept the queries or instructions in a form that the user enters and translate them into working instructions for the rest of the system. It also has to be able to translate the answers, produced by the system, into a form that the user can understand. Careful attention should be given to the screen design in order to make the expert system appear 'friendly' to the user.

(2.) *The Knowledge Base*
The knowledge base stores all the facts and rules about a particular problem domain. It makes these available to the inference engine in a form that it can use. The facts may be in the form of background information built into the system or facts that are input by the user during a consultation. The rules include both the production rules that apply to the domain of the expert system and the heuristics or rules-of-thumb that are provided by the domain expert in order to make the system find solutions more efficiently by taking short cuts.
(3.) The Shell or Inference Engine
The inference engine is the program that locates the appropriate knowledge in the knowledge base, and infers new knowledge by applying logical processing and problem-solving strategies.



Basic architecture of an expert system

b. What are the features of biological neural network that make it superior to even the most sophisticated AI computer system?

Answer: Refer Page No 213

Q.8 a. What are the advantages and disadvantages of neural networks?

Answer:
Advantages:

- A neural network can perform tasks that a linear program cannot.
- When an element of the neural network fails, it can continue without any problem by their parallel nature.
- A neural network learns and does not need to be reprogrammed.
- It can be implemented in any application and without any problem.
- Does not use pre-programmed knowledge base
- Suited to analyze complex pattern
- Have no restrictive assumptions
- Allows for qualitative data
- Can handle noisy data
- Can overcome autocorrelation
- User-friendly: clear output, and robust and flexible

Disadvantages:

- The neural network needs training to operate.
- The architecture of a neural network is different from the architecture of microprocessors therefore needs to be emulated.
- Requires high processing time for large neural networks.
- The neural network requires high quality data,
- Variables must be carefully selected a priori,
- Risk of overfitting,
- Requires a definition of architecture,
- Long processing time,
- Possibility of illogical network behavior, and

Large training sample required

b. Write comparison between neural networks and expert system.

Answer:

CHARACTERISTICS	**TRADITIONAL COMPUTING (including Expert Systems)**	**ARTIFICIAL NEURAL NETWORKS**
Processing style Functions	Sequential Logically (left brained) via Rules Concepts Calculations	Parallel Gestault (right brained) via Images Pictures Controls
Learning Method Applications	by rules (didactically) Accounting word processing math inventory digital communications	by example (Socratically) Sensor processing speech recognition pattern recognition text recognition

Table 1 Comparison of Computing Approaches

Characteristics	**Von Neumann Architecture Used for Expert Systems**	**Artificial Neural Networks**
Processors	VLSI (traditional processors)	Artificial Neural Networks; variety of technologies; hardware development is on going
Processing Approach	Separate	The same
Processing Approach	Processes problem rule at a one time; sequential	Multiple, simultaneously
Connections	Externally programmable	Dynamically self programming
Self learning	Only algorithmic parameters modified	Continuously adaptable
Fault tolerance	None without special processors	Significant in the very nature of the interconnected neurons
Neurobiology in design	None	Moderate

Programming	Through a rule based complicated	Self-programming; but network must be set up properly
Ability to be fast	Requires big processors	Requires multiple custom-built chips

Table 2 Comparisons of Expert Systems and Neural Networks.

Q.9 a. Explain role of AI in e-tourism.

Answer: Refer Page No 272

b. There are different types of clinical task to which expert system can be applied. Explain any such four tasks.

Answer: Refer Page No 286

<u>Text Book</u>

Introduction to Artificial Intelligence, Rajendra A Kerkar, PHI, 2005